KinBehR: KINect for Human BEHaviour Recognition

Cantarero, R., Santofimia-Romero, M., Lopez, J. C., Martinez-del-Rincon, J., & Nebel, J-C. (2014). *KinBehR: KINect for Human BEHaviour Recognition*. 89. Abstract from Web Intelligence Congress (WIC 2014), Warsaw, Poland. http://wic2014.mimuw.edu.pl/sites/default/files/Schedule_WIC2014.pdf

KinBehR: KINect for Human BEHaviour Recognition

Ruben Cantarero[1], Maria J. Santofimia[1], Juan Carlos Lopez[1],
Jesus Martinez-del-Rincon[2], and J.C. Nebel[3]

[1] University of Castilla-La Mancha, Spain
[2] Queen's University Belfast, United Kingdom
[3] Kingston University, United Kingdom
mariajose.santofimia@uclm.es

Key words: Human Action Recognition, Machine Learning, bag of words, Kinect.

Society is currently experimenting major changes in which, everyday more, major responsibilities are being delegated to technology to simplify people daily lives. That is the reason why, exponentially, systems and devices are targeting their role as people helpers. In this context, recognising human actions (Human Action Recognition) is playing a key role in important areas such as video surveillance, assistance to elder and disabled people, the Human-Computer Interaction, etc.

This work describes the implementation of a system for human action recognition, based on the novel use of depth cameras. This work therefore involves the evaluation of the use of depth cameras and how can it help improving recognition rates obtained for human actions in which body movements have not been previously showed to the recognising system. The followed methodology is decomposed into two stages. During first stage we have studied how to obtain and interpret the information provided by these cameras as well as to identify the information brought into light by the recognised actions. During the second stage, a machine learning algorithm has been implemented, using Bag of Words model. This model provides an estimate of the action classification from video sequences by analysing the image characteristics.

The use of different datasets, one for training and a different one for testing, enriches the system with robustness and multimodality support. In order to enhance the system with these features, a new dataset has been designed and recorded, also made available for the scientific community. Finally, obtained results have been compared with state-of-the-art studies to verify whether the proposed approach improves existing approaches.

The system evaluation and testing was carried out using a self-designed dataset known as KinbehrDataset, recorded using the module implemented in this work for image capture. To this end we had to prepare an appropriate context in which actors could naturally perform the actions the system had been trained with. Also, for future work, actors were also recording performing actions in a guided manner attending to provided commands. On the other hand, the unguided performance of actors were only based on the sole command of being themselves and staying in the room for 8 minutes. After having recorded all the videos, a process of segmentation and labeling was carried out. Finally, the system had to be tested using these videos and accuracy results were collected based on the ground truth information provided during the segmentation and labeling process.